A Comparative Study Between Wavelet Decomposition Method and Priestley's Method for Analyzing Real World Time Series

Lineesh M C*
and
C Jessy John†
Department of Mathematics,
National Institute of Technology Calicut
NIT Campus P O- 673 601
India.

September 14, 2009

Abstract

The increased computational more speed and developments in the area of algorithms have created the possibility to efficiently identify a well-fitting time series model for the given time series data. In this paper a new method is used for analyzing a given nonstationary/nonlinear time series data. The method is based on the wavelet decomposition technique where an algorithm is proposed to decompose the given time series in to trend series and detail series and then estimate each part separately using the well known threshold autoregressive model. Since the power spectral density can be used as a tool for identifying the most suitable model representing the time series a new method for estimating the power spectrum in the spectral domain using wavelet decomposition technique is introduced. The estimated power spectrum is used for finding the threshold value λ. The method is applied to real world time series data and the error comparison is carried out. In this paper a new method is used for analyzing a given

*E-mail:lineesh@nitc.ac.in
†E-mail:jessy@nitc.ac.in

nonstationary/nonlinear time series data. The method is based on the wavelet decomposition technique where an algorithm is proposed to decompose the given time series in to trend series and detail series and then estimate each part separately. As a comparative study real world time series data are analyzed using the developed method and also by using the Priestley's method and the error comparison is carried out.It is observed that the proposed method yields better results than the traditional method due to Priestley on account of its simplicity and less computational error. Then efficiency of the developed method illustrative examples like (i)Wolfer sunspot time series (ii) Laser time series and (iii) ECG time series are analyzed here. Also the adequacy of the model obtained in each case are verified using the normal probability plot.The method discussed here is an improvement over the conventional Box-Jenkins method on account of its simplicity and less computational time. Also the Mean Absolute Percentage Error (MAPE) are obtained in each case.

Keywords : Non Stationary Time Series; Wavelet Decomposition; Threshold Autoregressive Models; Scaling Coefficients; Wavelet Coefficients; Power Spectral Density; Periodogram.

1 Introduction

In time series analysis the covariance function of stochastic data obtained using the estimated models is used to compute the spectral density. The nonparametric method using windowed peridograms outperforms the conventional parametric method of estimating the PSD using autocovariances. Peridograms may be treated as squared Fourier transform of the given time series data. Time series models are to be preferred for spectral estimation if the true model is known to us. Hence the covariance estimates for higher lagged models is inefficient. Then the peridogram estimates are found to be more efficient as they are based on high lag covariance estimates.

According to the famous Wold's decomposition theorem a given time series can be decomposed in to a trend series and detail series which are orthogonal to each other. This idea gives the motivation to propose an algorithm to decomposes a given nonstationary/nonlinear time series into trend series and detail series.The trend series is analyzed using the ARMA models and the detail series is analyzed using the famous Threshold Autoregressive models estimated using the wavelet decomposition technique. The scale varying thresholds are estimated using the power spectrum estimated using wavelet decomposition of the given time series. In this paper real world time series data are analyzed using the developed method and a comparative study is made with the Priestley's method.

2 Wavelet Decomposition Method

An algorithm is developed here to analyze the given nonstationary/nonlinear time series and it is based on the concept of wavelet decomposition.
Given a nonstationary-nonlinear time series (Z_t), it is decomposed to a trend series $C_{M,t}$ and detail series

$$d_{M,t} = \sum_{j=1}^{M} d_{j,t} \tag{1}$$

where $d_{j,t}$ is the j^{th} level detail series, using wavelet decomposition technique. Then the trend series $C_{M,t}$ and detail series $d_{M,t}$ are analyzed separately. The model representing trend series can be obtained using any of the existing trend analysis method. The threshold autoregressive models TAR(k) are used to represent the detail series and the wavelet technique is applied to estimate the parameters of the TAR(k) model.

3 Methods for Power Spectrum Estimation

(1) Non-parametric Method

In this method the power spectrum is estimated as the discrete Fourier transform of the absolutely summable autocovarinace sequence.

(2) Parametric Method

In this method the power spectrum is estimated using the transfer function of the estimated model representing the time series. This method can provide a better frequency resolution than non-parametric method.

4 Power spectrum Estimation using Wavelet Transform

The discrete wavelet transform coefficients of the periodogram can be used for estimating the power spectrum.

Let $\{Z_t : t = 0, 1, 2, .., N-1\}$ be the given time series which has a defined logarithm of power spectral density $\ln G_Z(\exp(2\pi i f)), |f| \leq 0.5$.
Then logarithm of the periodogram $\ln P_Z[k]$ can be written as,

$$\ln P_Z[k] = \ln G_Z(e^{2\pi i f}) + \xi(e^{2\pi i f}) + \gamma \tag{2}$$

where $\xi(e^{2\pi i f})$ is a random process with probability distribution χ_2^2 with 2 degrees of freedom and $\gamma \approx 0.57721$ is the Euler-Mascheroni constant [2, 9].
The coefficients of the discrete wavelet transform with discrete time $\{C_{j,m}[k]\}$ are defined by the relation

$$C_{j,m}[k] = \sum_{k=0}^{N-1} (\ln P_Z[k] - \gamma)\psi_{j,m}[k]. \tag{3}$$

where $P_Z[k]$ are samples of the periodogram of the given time series of length $2N = 2^{M+1}$ obtained by the discrete Fourier transform.
Then the required estimate of the power spectral density is given by the inverse discrete wavelet transform,

$$\ln P_Z[k] = \frac{1}{N} \sum_{m=1}^{M-1} \sum_{j=1}^{2^m} C_{j,m}[k]\psi_{j,m}[k], \tag{4}$$

where $k = 0, 1, 2, ..., N-1$. Also the scale dependent thresholds λ_j for each level j is given by, $\lambda_j = C_{j,m} \ln N$.

5 Estimation of Threshold Autoregressive Model Using Wavelet Techniques

Let $\{Z_t : t = 0, 1, 2, ..., N-1\}$ be the given time series. Decomposing Z_t up to level M gives

$$Z_t = X_t + Y_t \tag{5}$$

where X_t is the trend series and Y_t is the detail series given by;

$$X_t = C_{M,t}; Y_t = d_{1,t} + d_{2,t} + ... + d_{M,t} : t = 0, 1, 2, ..., N-1;\ N = 2^J \tag{6}$$

where J is a positive integer.
Choose an appropriate scaling function ϕ and a wavelet function ψ (for example Haar) for the analysis.
Define the scaling and wavelet coefficients as follows;

$$\begin{aligned} \phi_{j,k}(t) &= 2^{-j/2}\phi(2^{-j}t - k) \quad j = 1, 2, 3, ..., J; k = 0, 1, 2, ..., 2^j - 1 \\ \psi_{j,k}(t) &= 2^{-j/2}\psi(2^{-j}t - k) \quad j = 1, 2, 3, ..., J; k = 0, 1, 2, ..., 2^j - 1 \end{aligned} \tag{7}$$

where

$$\phi(t) = \begin{cases} -2^{(-j/2)} & if \quad 2^j.k \le t < 2^j(k+1/2) \\ 2^{(-j/2)} & if \quad 2^j(k+1/2) \le t < 2^j(k+1) \end{cases} \tag{8}$$

and

$$\psi(t) = \begin{cases} 2^{(-j/2)} & if \quad 2^j.k \le t < 2^j(k+1/2) \\ -2^{(-j/2)} & if \quad 2^j(k+1/2) \le t < 2^j(k+1) \end{cases} \tag{9}$$

Define

$$\alpha_{j,k} = \sum_{t=0}^{N-1} \phi_{j,k}(t)X_t \tag{10}$$

and

$$\beta_{j,k} = \sum_{t=0}^{N-1} \psi_{j,k}(t)Y_t \tag{11}$$

Using (7),

$$X_t = \sum_{k=0}^{2^j-1} \alpha_{J,k}.\phi_{J,k}(t). \tag{12}$$

and using (8),

$$Y_t = \sum_{j=1}^{J} \sum_{k=0}^{2^j-1} \beta_{j,k}.\psi_{j,k}(t). \tag{13}$$

Under the wavelet decomposition method models representing trend series and detail series are separately estimated. The best fitting ARMA(p, q) or regression model, T(t), is considered for trend series.
The model representing the detail series is estimated using wavelet analysis as follows.

For $j = 1, 2, 3, ..., J$ define $\lambda_j = \sqrt{2log(\sharp d_{j,t})}$, where $(\sharp d_{j,t})$ denotes the cardinality of $\{d_{j,t}\}$. Also define $\lambda = \sqrt{2log(\sharp C_{N,t})}$.

The Threshold Autoregressive model representing the detail series $\{Y_t\}$ is given by,

$$Y_t = \begin{cases} b_1^{(1)}Y_{t-1} + b_2^{(1)}Y_{t-2} + ... + b_k^{(1)}Y_{t-k} + e_t^{(1)} & if \quad Y_{t-d} < \lambda \\ b_1^{(2)}Y_{t-1} + b_2^{(2)}Y_{t-2} + ... + b_k^{(2)}Y_{t-k} + e_t^{(2)} & if \quad Y_{t-d} \ge \lambda \end{cases} \tag{14}$$

where the coefficients $b_i^{(j)}$ are defined by,

$$b_j^{(i)} = \begin{cases} \sum_j \sum_t d_{j,t}^{(1)}.\psi_{j,t}^{(1)} \\ \sum_j \sum_t d_{j,t}^{(2)}.\psi_{j,t}^{(2)} \end{cases} \tag{15}$$

where

$$d_{j,t}^{(1)} = d_{j,t} \; if \; d_{j,t} < \lambda_j \; and \; d_{j,t}^{(2)} = d_{j,t} \; if \; d_{j,t} \geq \lambda_j \tag{16}$$

$$\psi_{j,t}^{(1)} = \psi_{j,t} \; if \; d_{j,t} < \lambda_j \; and \; \psi_{j,t}^{(2)} = \psi_{j,t} \; if \; d_{j,t} \geq \lambda_j. \tag{17}$$

The model representing the given time seriesZ_t, using wavelet decomposition is obtained by combining the model representing the trend series and the TAR(k) model which represent the detail series.

$$Z_t = \begin{cases} T(t) + b_1^{(1)} Y_{t-1} + b_2^{(1)} Y_{t-2} + ... + b_k^{(1)} Y_{t-k} + e_t^{(1)} & if \quad Y_{t-d} < \lambda \\ T(t) + b_1^{(2)} Y_{t-1} + b_2^{(2)} Y_{t-2} + ... + b_k^{(2)} Y_{t-k} + e_t^{(2)} & if \quad Y_{t-d} \geq \lambda \end{cases} \tag{18}$$

6 Analysis of Real World Time Series and PSD Estimation

(a) Sunspot Time Series

The annual sunspot numbers for the years 1700 - 1955 is considered for our analysis. The plot of the data is shown in figure 1.

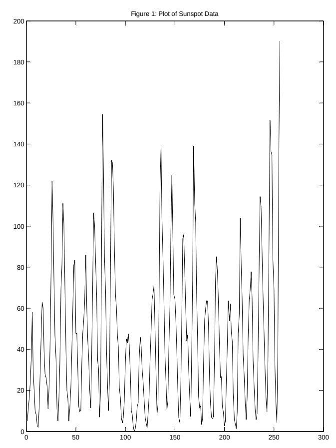

Figure 1: plot of sunspot data

The algorithm given above is used to analyze the sunspot time series data. The time series model is estimated using the developed method and the error analysis is done for checking the efficiency of the developed method. The details of the analysis are given in table 1.

Table 1: Analysis of Sunspot data using the Developed method

Threshold	Estimated Model	MAPE1	MSE 1
3.3302	$X_t = \begin{cases} 0.99X_{t-1} - 0.0028X_{t-2} \\ -0.62Y_{t-1} - 8.89Y_{t-2} \\ -21.7Y_{t-3} + e_t^{(1)} \quad if \quad Y_{t-1} < 3.33 \\ \\ 0.99X_{t-1} - 0.0028X_{t-2} \\ +0.0Y_{t-1} + 7.09Y_{t-2} \\ +5.484Y_{t-3} + e_t^{(2)} \quad if \quad Y_{t-1} \geq 3.33 \end{cases}$	0.7901%	4.4814

Model Estimation using Priestley's Method

A comparative study between the model estimated using the method of wavelet decomposition and the model estimated using Priestley's method[6] is done for examining the performance of the developed method.The analysis results of the sunspot data using Priestley's method is included in table 2. Also the normal probability plot of the residuals obtained using Priestley's method is given figure 3. The error comparison shows that the wavelet decomposition method is an improvement over the existing method due to Priestley.

Table 2: Analysis of Sunspot data using Priestley's method

Threshold	Estimated Model	MAPE2	MSE 2
35	$X_t = \begin{cases} 0.539X_{t-1} - 0.196X_{t-2} + 0.483X_{t-3} + e_t^{(1)} & if \quad X_{t-2} < 35 \\ 0.542X_{t-1} - 0.127X_{t-2} + 0.0168X_{t-3} + 0.051X_{t-4} + 0.029X_{t-5} + 0.45X_{t-6} + e_t^{(2)} & if \quad X_{t-2} \geq 35 \end{cases}$	3.1828%	6.5317

(b) Analysis of Stock Market Data

The data represents the monthly weighted-average exchange value of U. S. Dollar starting from September 1977 to December 1998. This is a secondary data[12]. The plot of the data is given in figure 4.

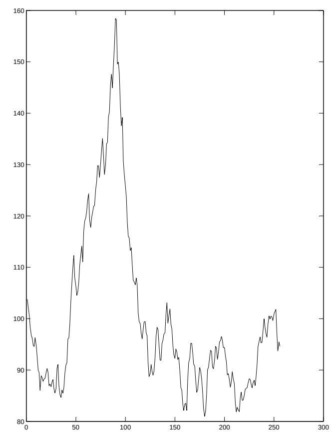

Figure 2: plot of stock exchange data

The stock market data is analyzed using the developed method. The time series model for the data is estimated and the details of analysis is given in table 3.

Table 3: Analysis of stock market data using the Developed method

Threshold	Estimated Model	MAPE1	MSE 1
56.3214	$X_t = \begin{cases} 0.99X_{t-1} - 0.0000778X_{t-2} \\ +0.1678Y_{t-1} + 1.2415Y_{t-2} \\ +e_t^{(1)} & if \quad Y_{t-1} < 56.321 \\ 0.99X_{t-1} - 0.000077X_{t-2} \\ +0.0008Y_{t-1} + 0.0Y_{t-2} \\ +e_t^{(2)} & if \quad Y_{t-1} \geq 56.321 \end{cases}$	1.12%	0.48

Model Estimation using Priestley's Method

The details of the analysis of the stock exchange data using Priestley's method is given in table 4. Also the normal probability plot of the residuals obtained using Priestley's method is given figure 6. The error comparison shows that the wavelet decomposition method is an improvement over the existing method due to Priestley.

Table 4: Analysis of Stock exchange data using Priestley's method

Threshold	Estimated Model	MAPE2	MSE 2
90	$X_t = \begin{cases} 1.003X_{t-1} - 0.4408X_{t-2} \\ +0.256X_{t-3} - 0.19X_{t-4} \\ +0.31X_{t-5} - 0.316X_{t-6} \\ +0.23X_{t-7} - 0.179X_{t-8} \\ +0.379X_{t-9} - 0.49X_{t-10} \\ +0.587X_{t-11} - 0.284X_{t-12} \\ +0.404X_{t-13} - 0.268X_{t-14} + e_t^{(1)} & if \quad X_{t-3} < 90 \\ \\ 1.169X_{t-1} - 0.1689X_{t-2} \\ +e_t^{(2)} & if \quad X_{t-3} \geq 90 \end{cases}$	1.9164%	0.5215

(d) Analysis of IBM Stock Price Data

The data represents the daily closing IBM stock prices[13]. The plot of the data is given in figure 10.

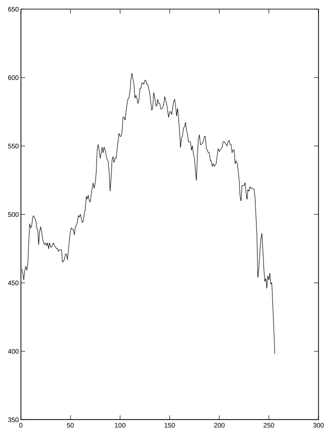

Figure 3: plot of IBM stock data

The IBM stock price data is analyzed using the developed method. The time series model for the data is estimated and the details of error analysis is given in table 7.

Table 5: Analysis of IBM data using the Developed method

Threshold	Estimated Model	MAPE1	MSE 1
3.11	$X_t = \begin{cases} 0.99X_{t-1} + 0.00023X_{t-2} \\ +0.31Y_{t-1} - 0.726Y_{t-2} \\ -8Y_{t-3} - 37.3Y_{t-4} + e_t^{(1)} & if \quad Y_{t-1} < 3.11 \\ 0.99X_{t-1} + 0.00023X_{t-2} \\ +0.0Y_{t-1} + 0.792Y_{t-2} \\ +1.98Y_{t-3} + 2.33Y_{t-4} + e_t^{(2)} & if \quad Y_{t-1} \geq 3.11 \end{cases}$	0.2044%	22.48

Model Estimation using Priestley's Method

The details of the analysis of the IBM stock data using Priestley's method is given in table 8. Also the normal probability plot of the residuals obtained using Priestley's method is given in figure 12. The error comparison shows that the wavelet decomposition method is an improvement over the existing method due to Priestley.

Table 6: Analysis of IBM Stock data using Priestley's method

Threshold	Estimated Model	MAPE2	MSE 2
560	$X_t = \begin{cases} 1.293X_{t-1} - 0.293X_{t-2} + e_t(1) & if \quad X_{t-1} < 560 \\ 1.13X_{t-1} - 0.338X_{t-2} + 0.17X_{t-3} + 0.145X_{t-4} - 0.28X_{t-5} + 0.016X_{t-6} - 0.106X_{t-7} + 0.257X_{t-8} + e_t^{(2)} & if \quad X_{t-1} \geq 560 \end{cases}$	3.665%	25.0007

7 Power Spectral Density Estimation

The following table summarizes the estimation results of the PSD of the real world time series discussed above using the wavelet decomposition method and the existing method due to Priestley.

Table 7: Estimation error due to PSD Estimation

Sr. No.	Time Series	Error(Wavelet Method)	Error(Priestley's Method)
1	Sun spot Time Series	0.45534	0.72438
2	Federal Stock Exchange data	0.50319	0.81323
3	IBM Stock price data	0.57085	0.65227

8 Conclusion

In this paper a new method for estimating the power spectrum of nonstationary-nonlinear time series using the wavelet decomposition is introduced and the method is compared with the existing method due to M. B. Priestley.

REFERENCES :

(1) Y.Kopsinis and S. McLaughlin,"Empirical Mode Decomposition Based Soft-Thresholding",Proceedings of the 16^{th} European Signal Processing Conference, EUSIPCO 2008.

(2) M. H. Hayes, Statistical Digital Signal Processing and Modeling, John Wiley and Sons, 2004.

(3) K. Ko, and M. Vannucci, Bayesian wavelet analysis of autoregressive fractionally integrated moving-average processes, Elsevier, Science Direct 136(2006), 3415-3434.

(4) A. Papoulis, Probability, Random Variables, and Stochastic Processes (3rd Edition), McGraw-Hill, 1991.

(5) D. B. Percival and A. T. Walden, Wavelet Methods for Time Series Analysis, Cambridge University Press, 2000.

(6) M. B. Priestley, Non-linear and Non-stationary Time Series Analysis,Academic Press, 1988.

(7) R. M. Rao and A. S. Bopardikar, Wavelet Transforms - Introduction to Theory and Applications, Pearson Education, 1998.

(8) S. Soltani, On the use of the wavelet decomposition for time series prediction,Elsevier, Neurocomputing 48 (2002) 267-277.

(9) W. W. S. Wei, Time Series Analysis-Univariate and Multivariate Methods, Addison-Wesley Publishing Company, 1990.

(10) S. Mallat, A wavelet tour of signal processing , Academic Press, 1999.

(11) Eugenia B. Christopouloul , Athanassios N. Skodras and Alexander A. Georgakilas, Time Series Analysis of Sunspot Oscillations Using the Wavelet Transform.

(12) www.FreeLunch.com-http://www.eonomy.com/freelunch.

(13) www.robhyndman.info.

****